I Do Not Know What You Visited Last Summer: Protecting users from stateful third-party web tracking with TrackingFree Browser

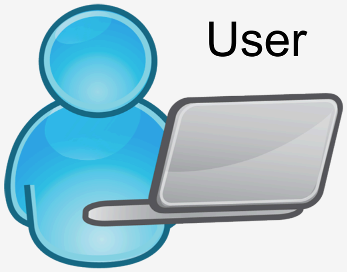
Xiang Pan[§], Yinzhi Cao[†], Yan Chen[§]

[§] Northwestern University
[†] Columbia University

Roadmap

- **Introduction & Background**
- System Design
- Evaluation
- Conclusion

Web Tracking



User

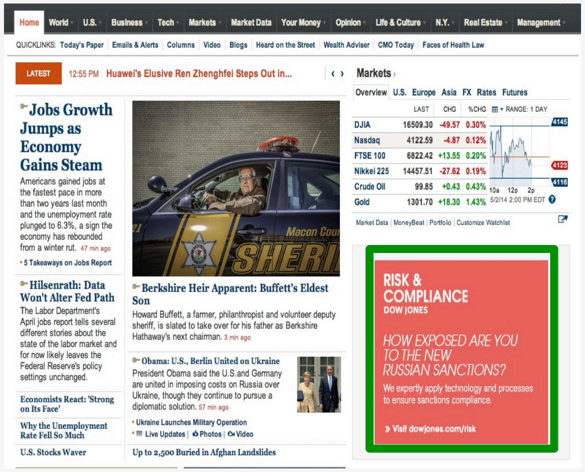
Referer : http://online.wsj.com/
Cookie : id = 12345

Referer : http://www.cnn.com/
Cookie : id = 12345

Tracking server

visit

visit



Home | World | U.S. | Business | Tech | Markets | Market Data | Your Money | Opinion | Life & Culture | N.Y. | Real Estate | Management

QUICKLINKS: Today's Paper | Emails & Alerts | Columns | Video | Blogs | Heard on the Street | Wealth Adviser | CMO Today | Faces of Health Law

LATEST 12:55 PM Huawei's Elusive Ren Zhenghfei Steps Out in...

Jobs Growth Jumps as Economy Gains Steam

Americans gained jobs at the fastest pace in more than two years last month and the unemployment rate plunged to 6.3%, a sign the economy has rebounded from a winter rut. 47 min ago

- 5 Takeaways on Jobs Report

Hilsenrath: Data Won't Alter Fed Path

The Labor Department's April jobs report tells several different stories about the state of the labor market and for now likely leaves the Federal Reserve's policy settings unchanged.

Economists React: 'Strong on Its Face'

Why the Unemployment Rate Fell So Much

U.S. Stocks Waver

Berkshire Heir Apparent: Buffett's Eldest Son

Howard Buffett, a farmer, philanthropist and volunteer deputy sheriff, is slated to take over for his father as Berkshire Hathaway's next chairman. 3 min ago

Obama: U.S., Berlin United on Ukraine

President Obama said the U.S.and Germany are united in imposing costs on Russia over Ukraine, though they continue to pursue a diplomatic solution. 57 min ago

- Ukraine Launches Military Operation
- Live Updates | Photos | Video

Up to 2,500 Buried in Afghan Landslides

Markets ›

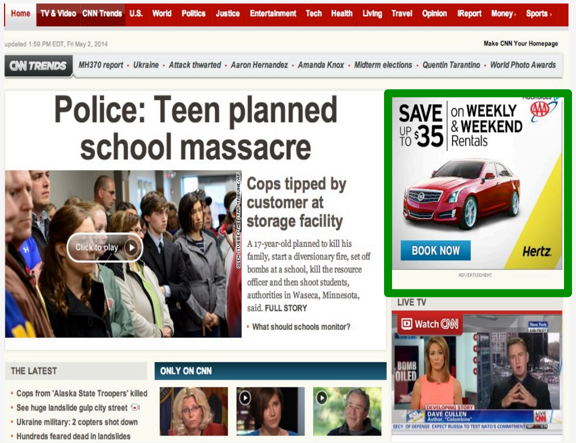
Overview | U.S. | Europe | Asia | FX | Rates | Futures

	LAST	CHG	%CHG
DJIA	16509.30	-49.57	0.30%
Nasdaq	4122.59	-4.87	0.12%
FTSE 100	6822.42	+13.55	0.20%
Nikkei 225	14457.51	-27.62	0.19%
Crude Oil	99.85	+0.43	0.43%
Gold	1301.70	+18.30	1.43%

RANGE: 1 DAY

5/2/14 2:00 PM EDT

Market Data | MoneyBeat | Portfolio | Customize Watchlist



Home | TV & Video | CNN Trends | U.S. | World | Politics | Justice | Entertainment | Tech | Health | Living | Travel | Opinion | iReport | Money › | Sports ›

updated 1:59 PM EDT, Fri May 2, 2014

Make CNN Your Homepage

CNN TRENDS MH370 report · Ukraine · Attack thwarted · Aaron Hernandez · Amanda Knox · Midterm elections · Quentin Tarantino · World Photo Awards

Police: Teen planned school massacre

Cops tipped by customer at storage facility

A 17-year-old planned to kill his family, start a diversionary fire, set off bombs at a school, kill the resource officer and then shoot students, authorities in Waseca, Minnesota, said. FULL STORY

- What should schools monitor?

GLEN STUBBE/THE STAR TRIBUNE/AP

THE LATEST

- Cops from 'Alaska State Troopers' killed
- See huge landslide gulp city street
- Ukraine military: 2 copters shot down
- Hundreds feared dead in landslides

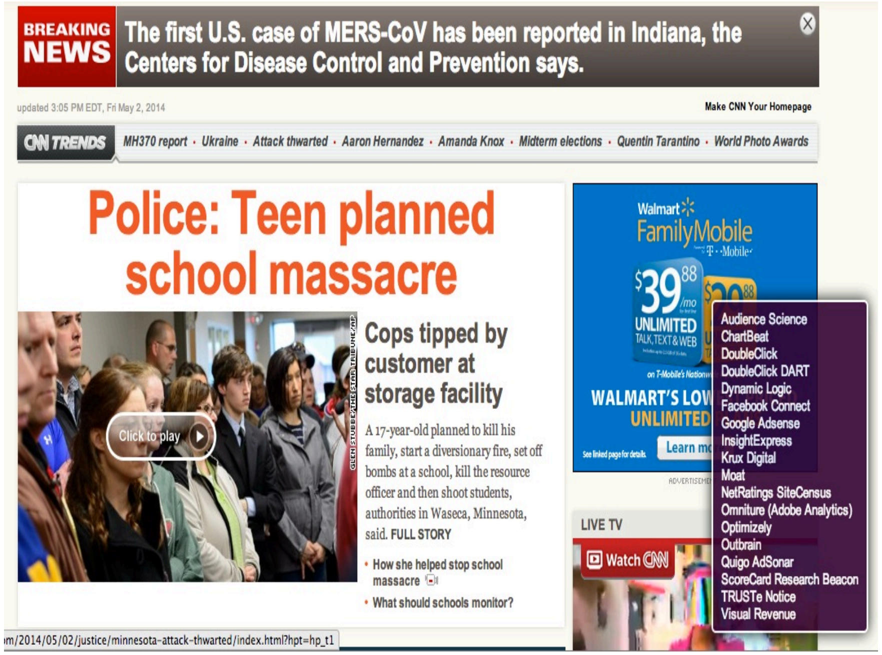
ONLY ON CNN

LIVE TV

Web Tracking is Prevalent and Serious

- Prevalent
  - More than 90% of Alexa Top 500 websites [Roesner, NSDI 2012].
  - A web page usually has multiple tracking elements.

Web Tracking is Prevalent and Serious



BREAKING NEWS The first U.S. case of MERS-CoV has been reported in Indiana, the Centers for Disease Control and Prevention says.

updated 3:05 PM EDT, Fri May 2, 2014

Make CNN Your Homepage

CNN TRENDS MH370 report · Ukraine · Attack thwarted · Aaron Hernandez · Amanda Knox · Midterm elections · Quentin Tarantino · World Photo Awards

Police: Teen planned school massacre

Cops tipped by customer at storage facility

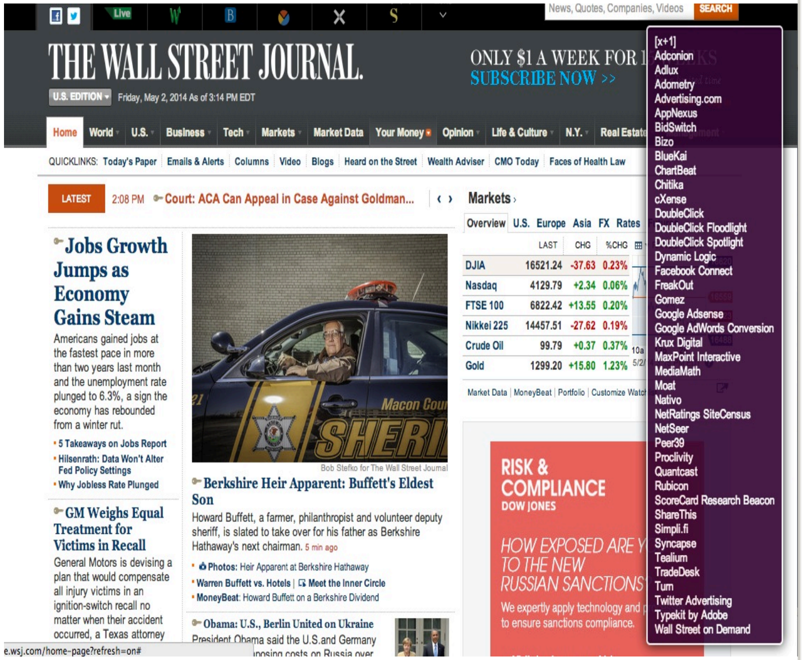
A 17-year-old planned to kill his family, start a diversionary fire, set off bombs at a school, kill the resource officer and then shoot students, authorities in Waseca, Minnesota, said. FULL STORY

- How she helped stop school massacre
- What should schools monitor?

Click to play
GLEN STUBBE/THE STAR TRIBUNE/AP
Walmart FamilyMobile T-Mobile
$39.88/mo
UNLIMITED TALK, TEXT & WEB
WALMART'S LOW
UNLIMITED
Learn mo
ADVERTISEMEN
LIVE TV
Watch CNN

Audience Science
ChartBeat
DoubleClick
DoubleClick DART
Dynamic Logic
Facebook Connect
Google Adsense
InsightExpress
Krux Digital
Moat
NetRatings SiteCensus
Omniture (Adobe Analytics)
Optimizely
Outbrain
Quigo AdSonar
ScoreCard Research Beacon
TRUSTe Notice
Visual Revenue

m/2014/05/02/justice/minnesota-attack-thwarted/index.html?hpt=hp_t1



THE WALL STREET JOURNAL.

U.S. EDITION Friday, May 2, 2014 As of 3:14 PM EDT

ONLY $1 A WEEK FOR 1
SUBSCRIBE NOW >>

Home · World · U.S. · Business · Tech · Markets · Market Data · Your Money · Opinion · Life & Culture · N.Y. · Real Estate

QUICKLINKS: Today's Paper | Emails & Alerts | Columns | Video | Blogs | Heard on the Street | Wealth Adviser | CMO Today | Faces of Health Law

LATEST 2:08 PM Court: ACA Can Appeal in Case Against Goldman...

Markets ›

Overview U.S. Europe Asia FX Rates

	LAST	CHG	%CHG
DJIA	16521.24	-37.63	0.23%
Nasdaq	4129.79	+2.34	0.06%
FTSE 100	6822.42	+13.55	0.20%
Nikkei 225	14457.51	-27.62	0.19%
Crude Oil	99.79	+0.37	0.37%
Gold	1299.20	+15.80	1.23%

Market Data | MoneyBeat | Portfolio | Customize Watch

Jobs Growth Jumps as Economy Gains Steam

Americans gained jobs at the fastest pace in more than two years last month and the unemployment rate plunged to 6.3%, a sign the economy has rebounded from a winter rut.

- 5 Takeaways on Jobs Report
- Hilsenrath: Data Won't Alter Fed Policy Settings
- Why Jobless Rate Plunged

GM Weighs Equal Treatment for Victims in Recall

General Motors is devising a plan that would compensate all injury victims in an ignition-switch recall no matter when their accident occurred, a Texas attorney

Bob Stefko for The Wall Street Journal

Berkshire Heir Apparent: Buffett's Eldest Son

Howard Buffett, a farmer, philanthropist and volunteer deputy sheriff, is slated to take over for his father as Berkshire Hathaway's next chairman. 5 min ago

- Photos: Heir Apparent at Berkshire Hathaway
- Warren Buffett vs. Hotels | Meet the Inner Circle
- MoneyBeat: Howard Buffett on a Berkshire Dividend

Obama: U.S., Berlin United on Ukraine

President Obama said the U.S.and Germany

Macon Cou
SHERI
RISK & COMPLIANCE
DOW JONES
HOW EXPOSED ARE Y
TO THE NEW RUSSIAN SANCTIONS
We expertly apply technology and
to ensure sanctions compliance.
10a
5/2/

[x+1]
Adconion
Adlux
Adometry
Advertising.com
AppNexus
BidSwitch
Bizo
BlueKai
ChartBeat
Chitika
cXense
DoubleClick
DoubleClick Floodlight
DoubleClick Spotlight
Dynamic Logic
Facebook Connect
FreakOut
Gomez
Google Adsense
Google AdWords Conversion
Krux Digital
MaxPoint Interactive
MediaMath
Moat
Nativo
NetRatings SiteCensus
NetSeer
Peer39
Proclivity
Quantcast
Rubicon
ScoreCard Research Beacon
ShareThis
Simpli.fi
Syncapse
Tealium
TradeDesk
Turn
Twitter Advertising
Typekit by Adobe
Wall Street on Demand

e.wsj.com/home-page?refresh=on#

Web Tracking is Prevalent and Serious

- Prevalent
    - More than 90% of Alexa Top 500 websites [Roesner, NSDI 2012].
    - A web page usually has multiple tracking elements.
- Serious
    - Not only browsing history, but also other sensitive information such as location, name and email, will be leaked out.
    - Potential for abuse is enormous.

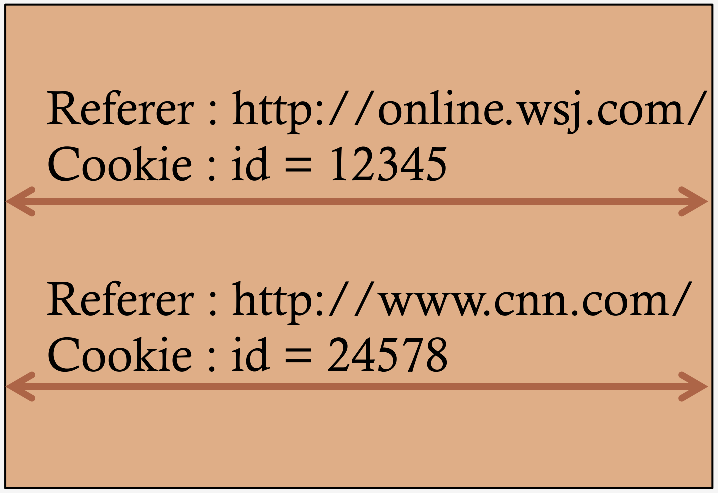
No Effective Defense Approach

- Disable third-party cookie
  - Many other storages to store user’s identifier.
- Blacklist-based anti-tracking tools
  - Priori knowledge of tracking servers.
- *Do-Not-Track* header
  - No enforcement.

TrackingFree

Goals and Challenges

- Complete Anti-tracking Capability
- Backward Compatibility
- Affordable Performance



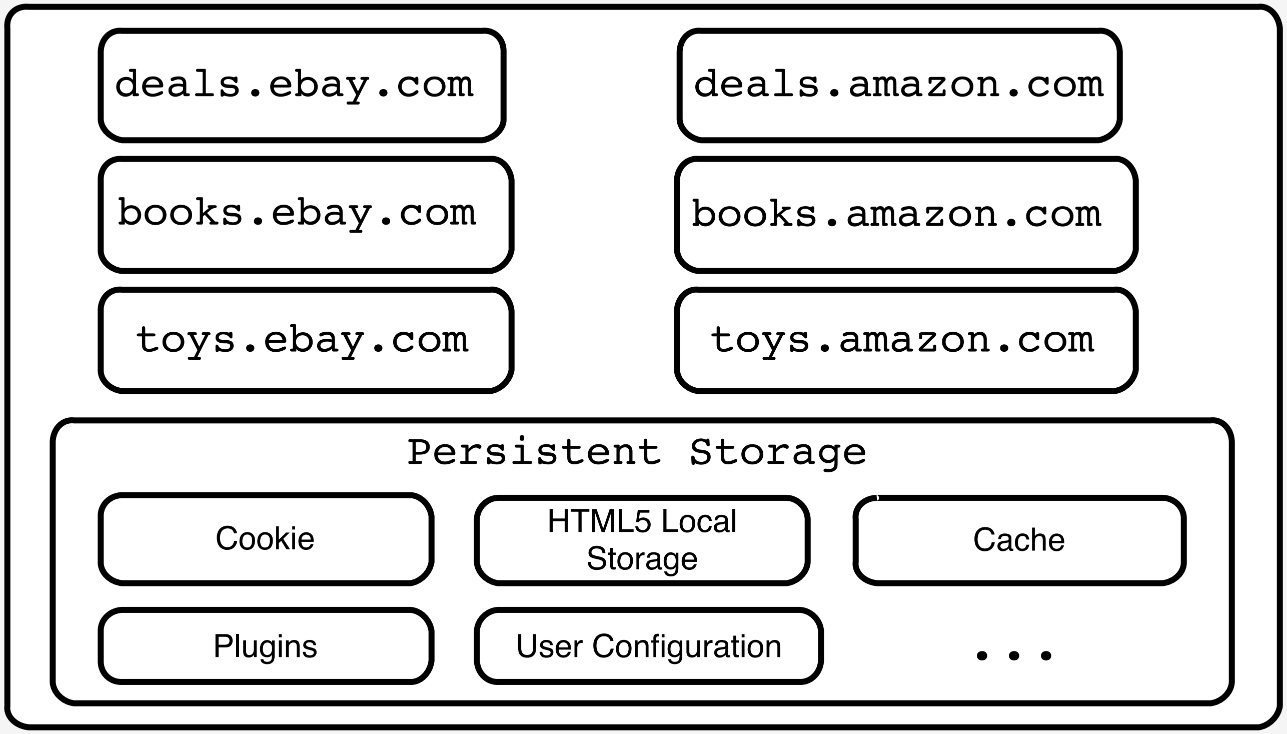
Referer : http://online.wsj.com/
Cookie : id = 12345

Referer : http://www.cnn.com/
Cookie : id = 24578

Core Idea : TrackingFree partitions client-side states into multiple isolation units so that the identifiers still exist but not unique any more!

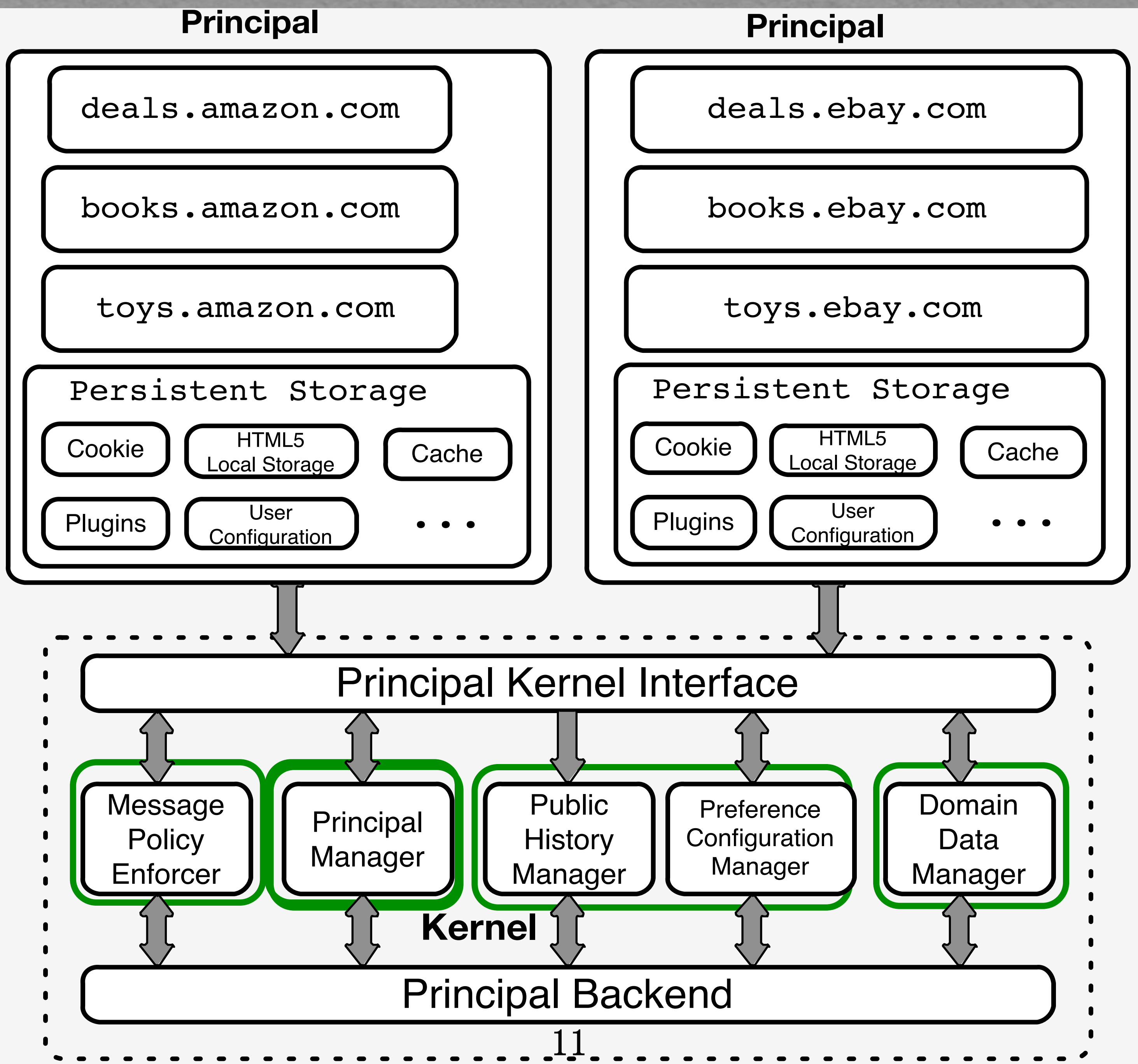
Roadmap

- Introduction & Background
- **System Design**
- Evaluation
- Conclusion



Regular Browser Architecture

TrackingFree Architecture



Principal

deals.amazon.com

books.amazon.com

toys.amazon.com

Persistent Storage

Cookie

HTML5 Local Storage

Cache

Plugins

User Configuration

. . .

Principal

deals.ebay.com

books.ebay.com

toys.ebay.com

Persistent Storage

Cookie

HTML5 Local Storage

Cache

Plugins

User Configuration

. . .

Principal Kernel Interface

Message Policy Enforcer

Principal Manager

Public History Manager

Preference Configuration Manager

Domain Data Manager

Kernel

Principal Backend

Contents Allocation Mechanism

- Initial Contents Allocation
  - Handles those top frames that are navigated by users directly.
  - Based on registered domain name (e.g. google.com, sina.com.cn).
- Derivative Contents Allocation
  - Handles those frames that are generated due to the contents on other frames, which we call child frame.

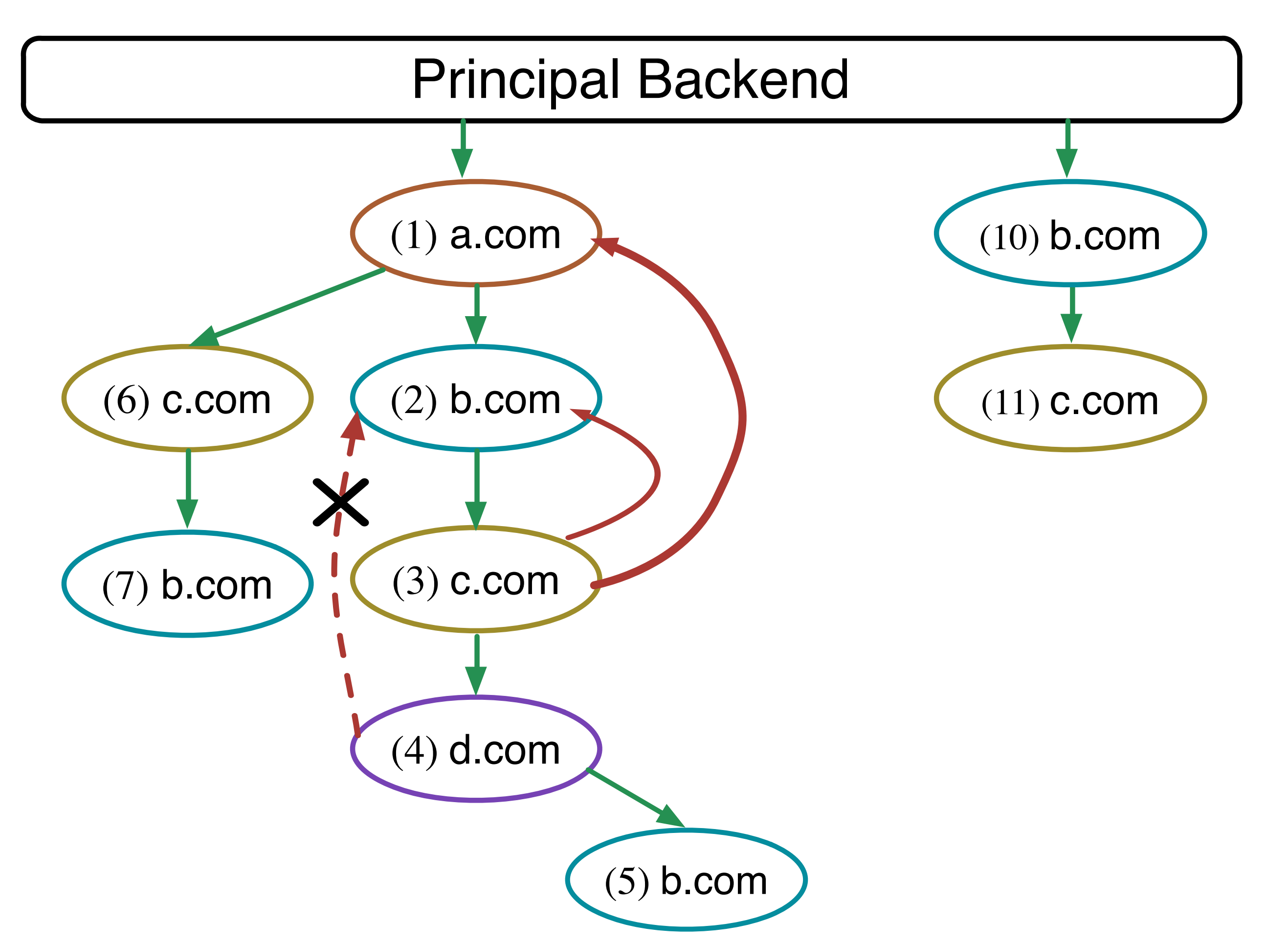
Derivative Contents Allocation

- Principal Switch
  - Cross-domain
  - User-triggered
- Principal Selection
  - Maintains an in-degree-bounded graph for principals.
  - The in-degree of the graph is set to two.

Derivative Contents Allocation

- F

- F



Roadmap

- Introduction & Background
- System Design
- **Evaluation**
- Conclusion

Evaluation

- Anti-tracking capability
  - Formal proof
  - Experiments with real world websites
- Compatibility
- Performance
  - Latency
  - Memory usage
  - Disk usage

Formal Proof

- Methodology
  - Use Alloy to formally analyze TrackingFree ’s anti-tracking ability.
  - Describe TrackingFree’s behaviors on an existing Alloy web model [Akhawe et al. CSF 2010].

- Results
  - Formally verified that trackers can correlate TrackingFree user’s activities up to three principals without site collaboration.

Anti-tracking Capability with Real World Websites

- Re-implemented an in-complete but accurate tracking token detection approach proposed on [Roesner et al. NSDI 2012].
- The approach is based on the observation that each tracking request must contain the user's globally unique identifier.

Anti-tracking Capability with Real World Web Sites

Tracking Host	Prevalence (# Domains)	Tracking Token(s)
b.scorecardresearch.com	133	UIDR
ad.doubleclick.net	117	id, __gads
ib.adnxs.com	75	anj
p.twitter.com	70	__utma
cm.g.doubleclick.net	56	id
ad.yieldmanager.com	52	bx
bs.serving-sys.com	40	A4
cdn.api.twitter.com	40	__utmz
secure-us.imrworldwide.com	38	IMRID
adfarm.mediaplex.com	31	svid

Top 10 Tracking Hosts

Anti-tracking Capability with Real World Web Sites

Tracking Host	Prevalence (# Domains)	Tracking Token(s)
b.scorecardresearch.com	133	UIDR
ad.doubleclick.net	117	id, __gads
ib.adnxs.com	75	anj
p.twitter.com	70	__utma
bs.serving-sys.com	40	A4
cdn.api.twitter.com	40	__utmz
secure-us.imrworldwide.com	38	IMRID
adfarm.mediaplex.com	31	svid

Top 10 Tracking Hosts

- TrackingFree eliminated all of them.

Compatibility

- Manually tested TrackingFree's compatibility on 69 third-party services from Alexa Top 50 websites.

Name	Example	# Succeeded Instance	# Total Instance
Cross-site online payment	Purchase on Ebay and make payment on Paypal	1	1
Cross-site content sharing	Share Youtube video to Facebook account	32	32
Signle sign-on	Using Facebook account to login Yahoo	35	36
Overall Results		68	69

Performance

Source	Overhead
Latency	~3%-~20%
Memory	~25MB/principal
Disk	~0.6MB/principal

Roadmap

- Introduction & Background
- System Design
- Evaluation
- **Conclusion**

Conclusion

- We designed and implemented TrackingFree browser which completely protects users from third-party web tracking by isolating web contents.
- We theoretically and experimentally proved TrackingFree's anti-tracking capability.
- TrackingFree is backward compatible with existing websites.

Thanks & Questions?

http://list.cs.northwestern.edu/WebSecurity

Backup slides …

Out-of-scope threats

- Within-site Tracking
- Tracking by exploiting browser vulnerabilities
- Stateless tracking

Preference Configuration Manager

- User preference can be abused to store tracking identifier. (e.g. strict transport security)
- Completely isolating user preference affects user preference.
- Our solution:
  - Isolate user preference.
  - Apply user-initiated changes to all of the principals.
  - Monitor GUI message to determine user-initiated preference change.

Related Work

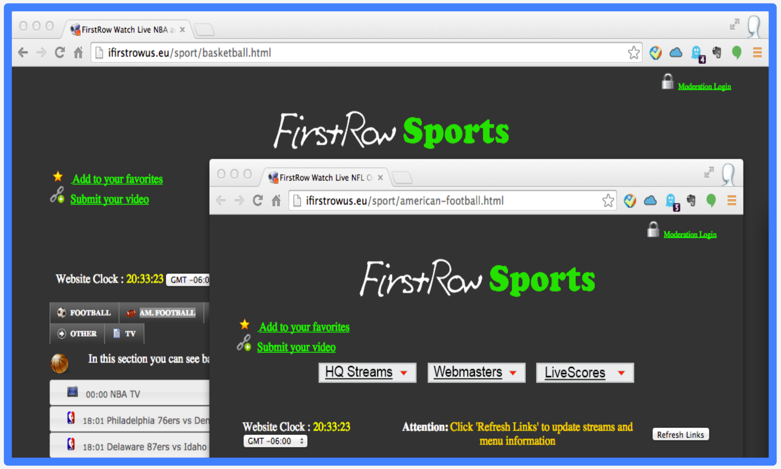
Browser	Isolation Mechanism	Contents Allocation Mechanism	Anti-tracking Capability
IE8	In-memory Isolation	Tab based	No
Chromium	In-memory Isolation	Top-frame based	No
Gazelle	In-memory Isolation	SOP based	No
OP	In-memory Isolation	Web Page based	No
AppIsolation	Technique-specific Storage	User Configuration based	Not complete
Tahoma	Virtual Machine	User Configuration based	Not complete
Stainless	Technique-specific Storage	User Configuration based	Not complete
Fluid, MultiFirefox	Profile	User Configuration based	Not complete
TrackingFree	Profile	Indegree-bounded Principal Graph based	**Complete**

Principal Switch

- Two intuitive yet extreme policies：
  - Not privacy-preserving (no switch)
  - Unnecessary overhead (switch all the time)
- Our solution: switch principal only if the following two conditions are met:
  - Cross-site
  - User-triggered

Principal Switch

Same principal

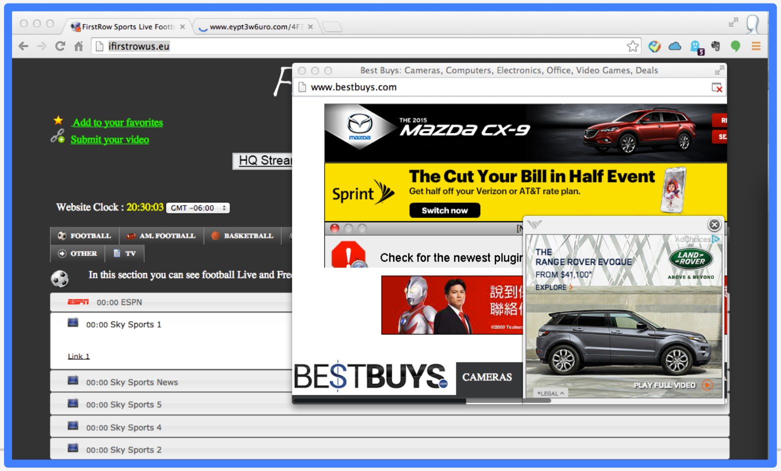


Website Clock : 20:33:23 GMT -06:0

In this section you can see ba

Website Clock : 20:33:23 GMT -06:00

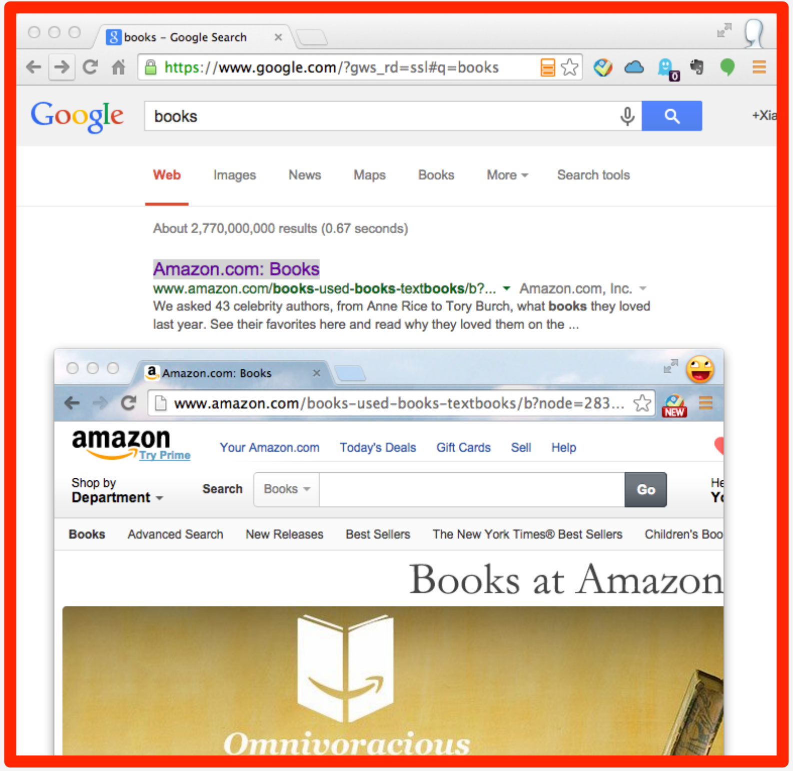
Attention: Click 'Refresh Links' to update streams and menu information



Website Clock : 20:30:03 GMT -06:00

In this section you can see football Live and Fre

Different principal



About 2,770,000,000 results (0.67 seconds)

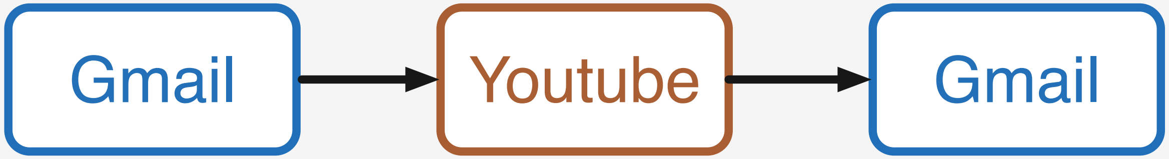
Amazon.com: Books
www.amazon.com/books-used-books-textbooks/b?... Amazon.com, Inc.
We asked 43 celebrity authors, from Anne Rice to Tory Burch, what **books** they loved last year. See their favorites here and read why they loved them on the ...

Principal Selection

- Two intuitive yet extreme policies:
  - Break compatibility (always create new principal)
  - Break anti-tracking capacity (create at most one principal for each domain)

Principal Selection

- Two intuitive yet extreme policies:
  - Break compatibility (always create new principal)
  - Break anti-tracking capacity (create at most one principal for each domain)

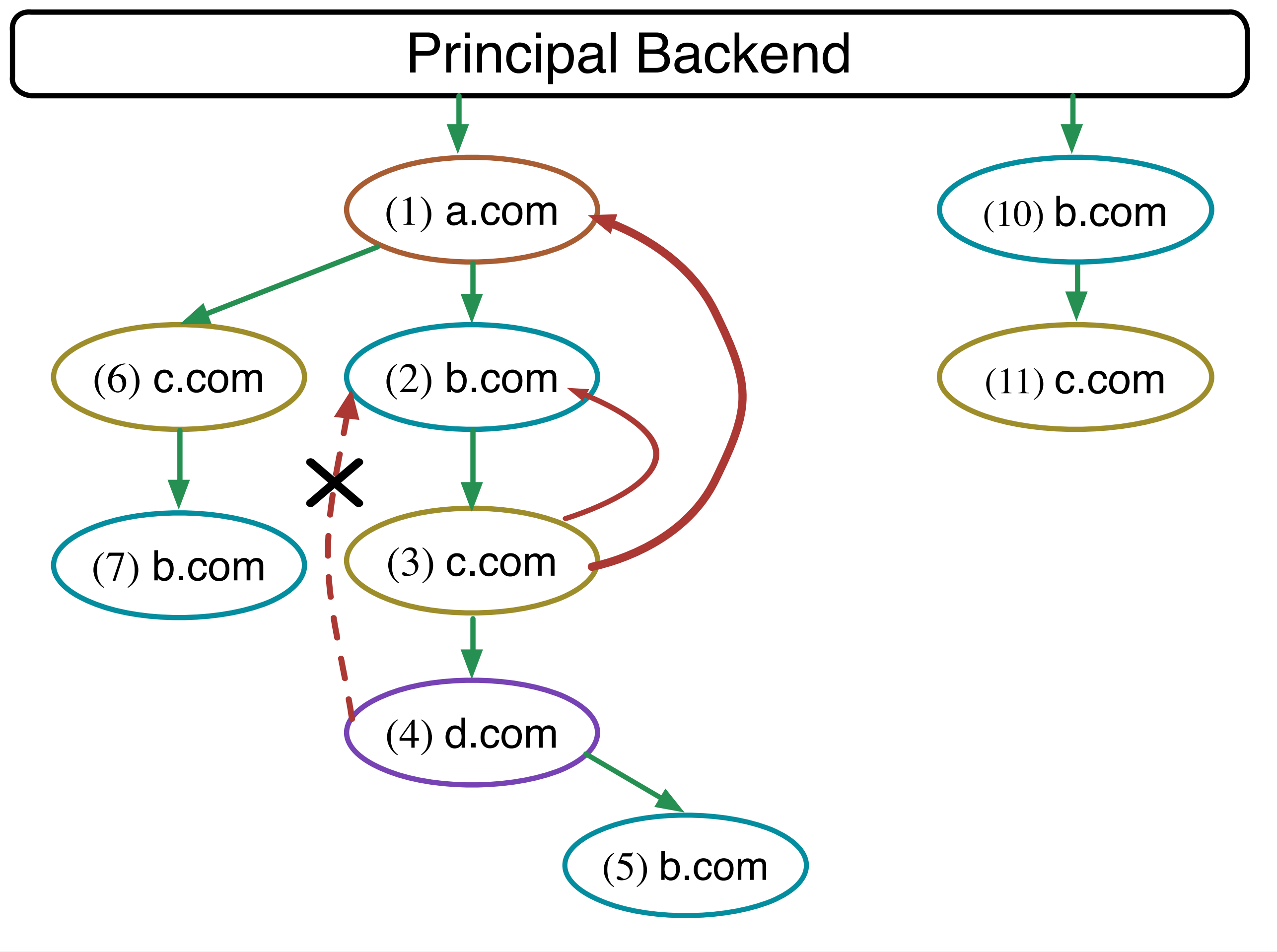


Principal Selection

- Two intuitive yet extreme policies:
  - Break compatibility (always create new principal)
  - Break anti-tracking capacity (create at most one principal for each domain)
- Our solution:
  - Maintains an in-degree-bounded graph for principals.
  - The in-degree of the graph is set to two.

Principal Selection

- T
  - pal)
  - ne
- C
  -
  -

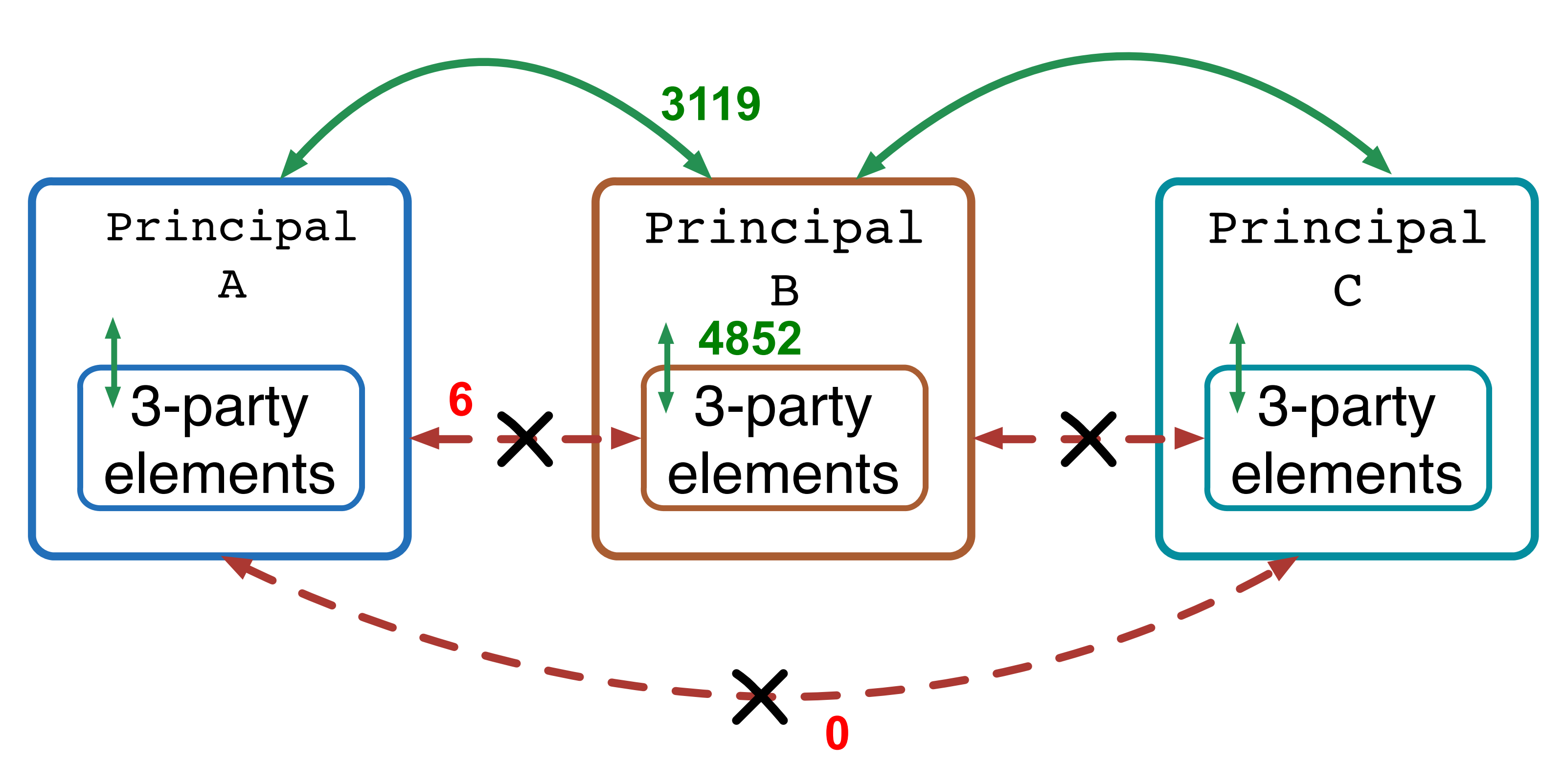


Principal Communication

- Explicit communication is widely used, but break the isolation mechanism.
- Our solution: we restrict the use of explicit communication as follows:
  - Third-party elements in one principle can not explicitly communicate with other principals.
  - First-party elements can only explicitly communicate with the first-party elements placed in its neighbor principals

Principal Communication

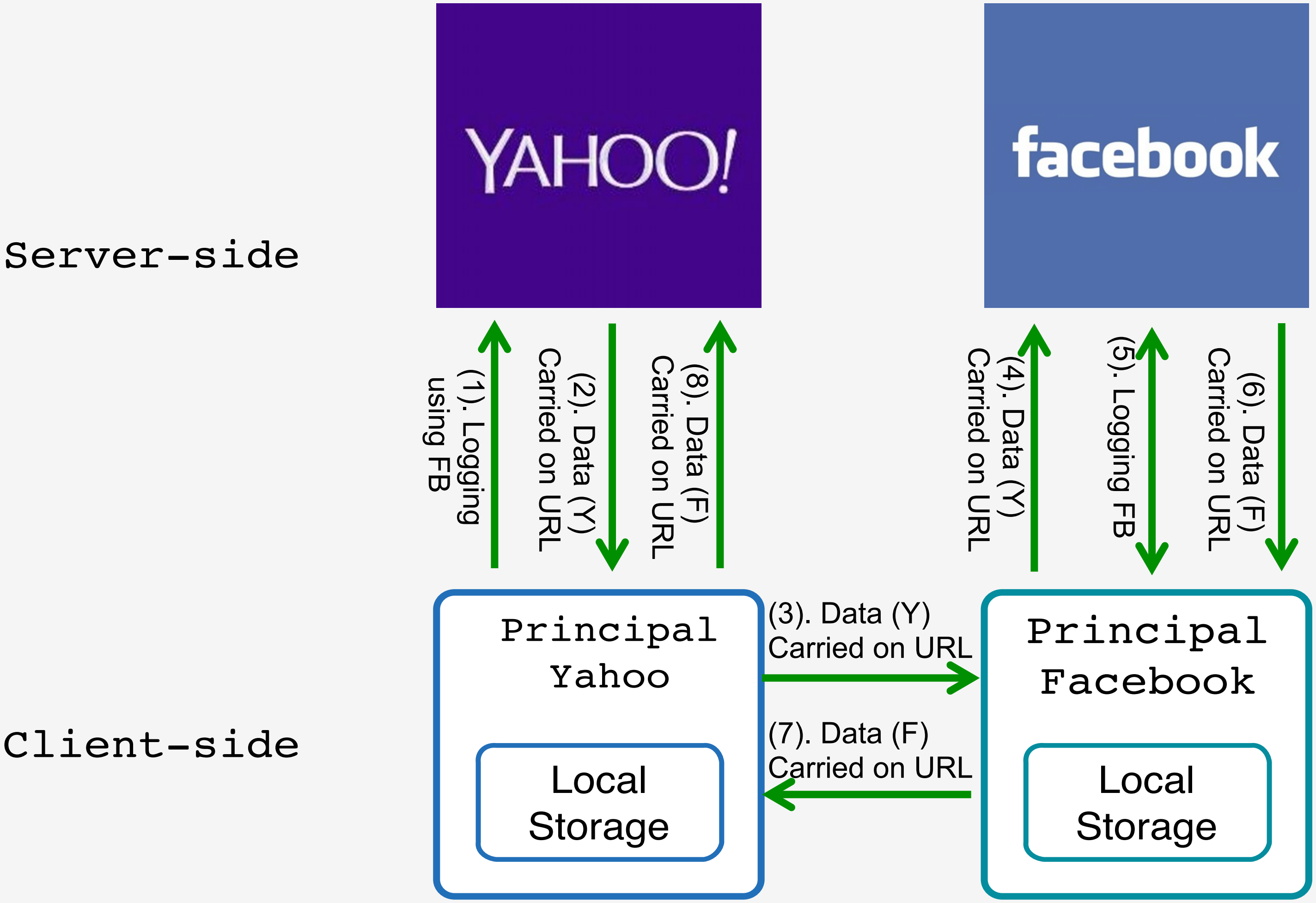
- Explicit communication is widely used, but break



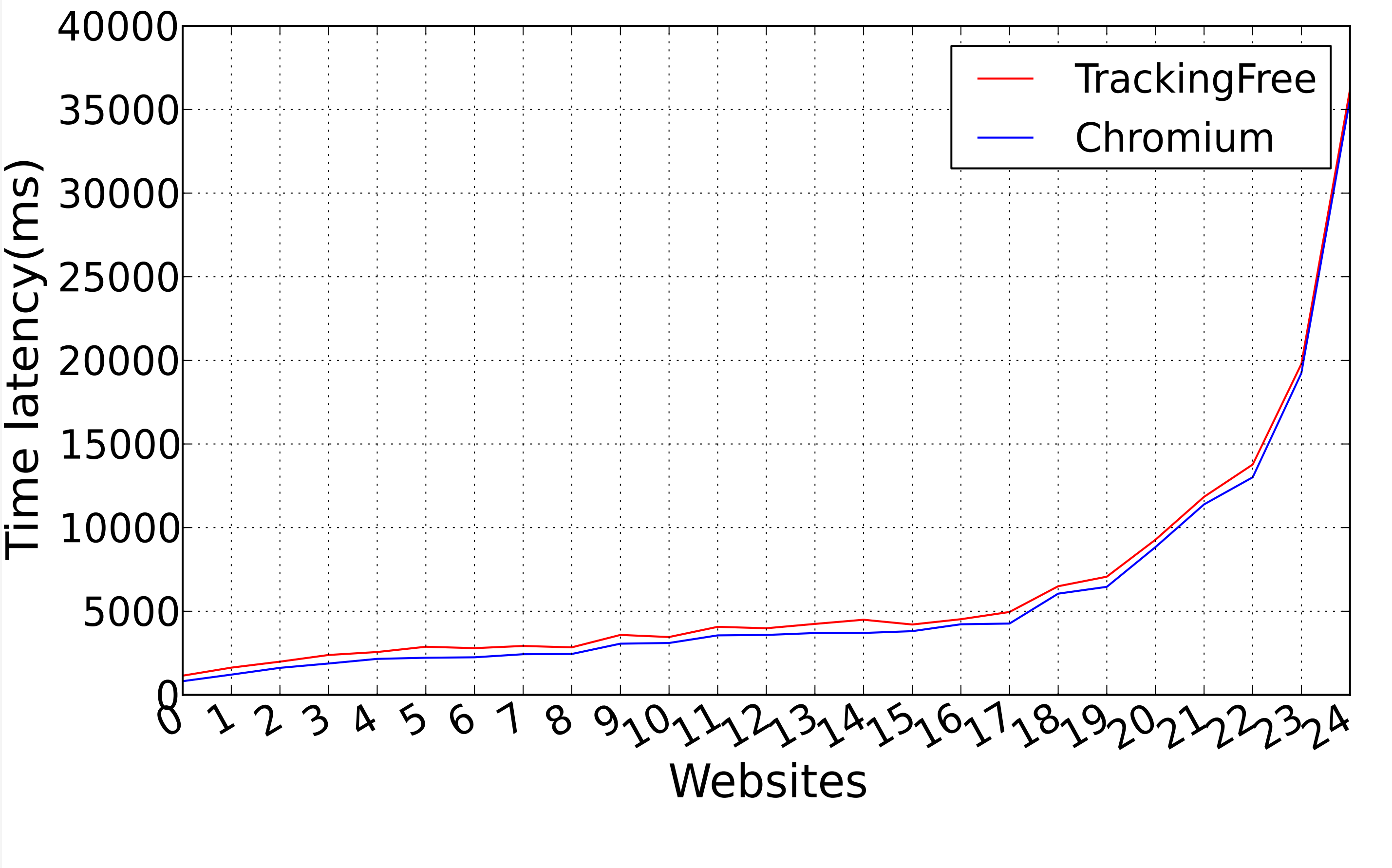
Principal Communication

- Implicit Communication
  - History sharing (e.g. history, bookmarks)
  - User preference sharing
  - Communication through navigation URL parameters

Case study: Logging Yahoo using Facebook Account

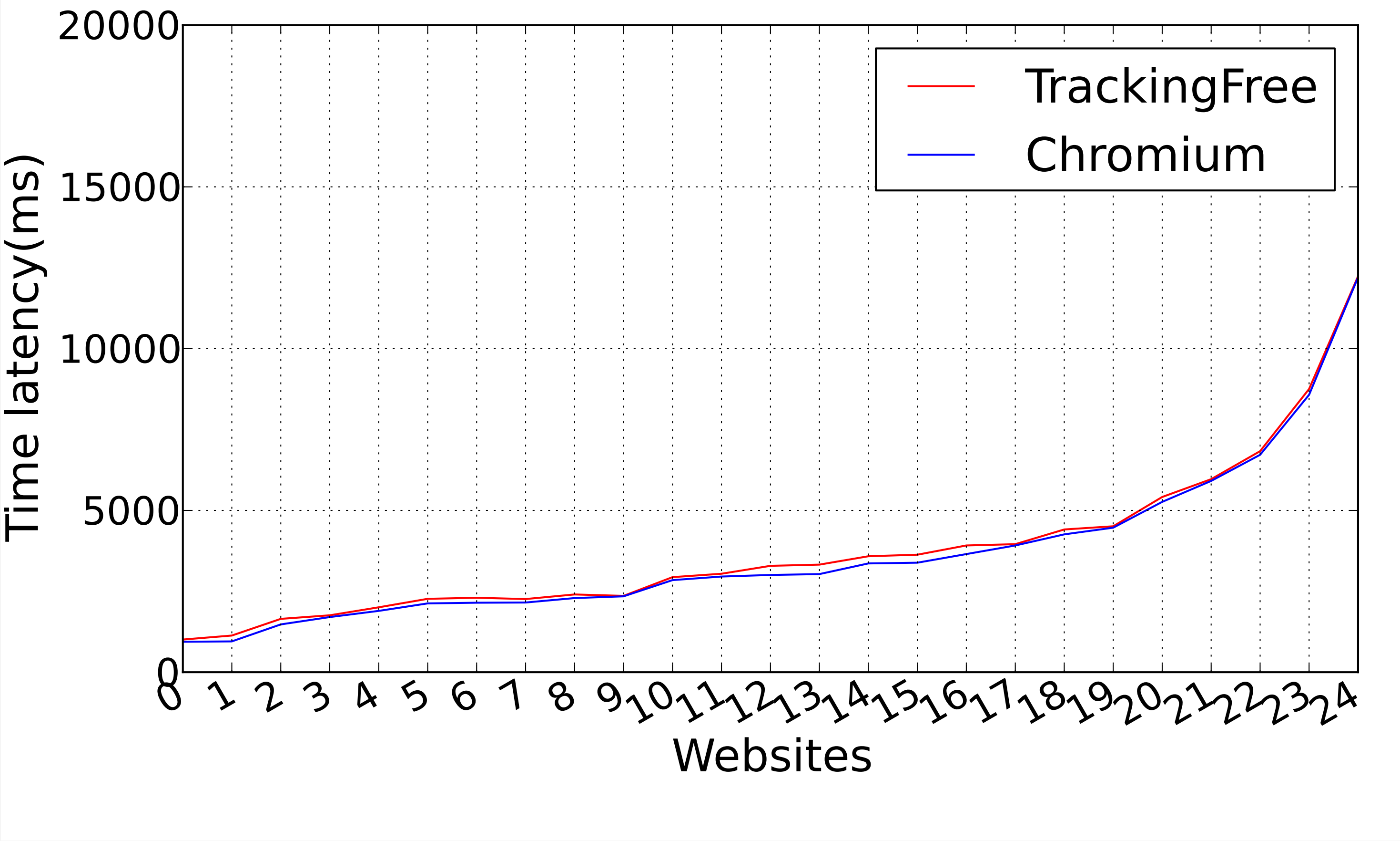

Server-side
YAHOO!
facebook
(1). Logging using FB
(2). Data (Y) Carried on URL
(8). Data (F) Carried on URL
(4). Data (Y) Carried on URL
(5). Logging FB
(6). Data (F) Carried on URL
Client-side
Principal Yahoo
Local Storage
(3). Data (Y) Carried on URL
(7). Data (F) Carried on URL
Principal Facebook
Local Storage

Performance



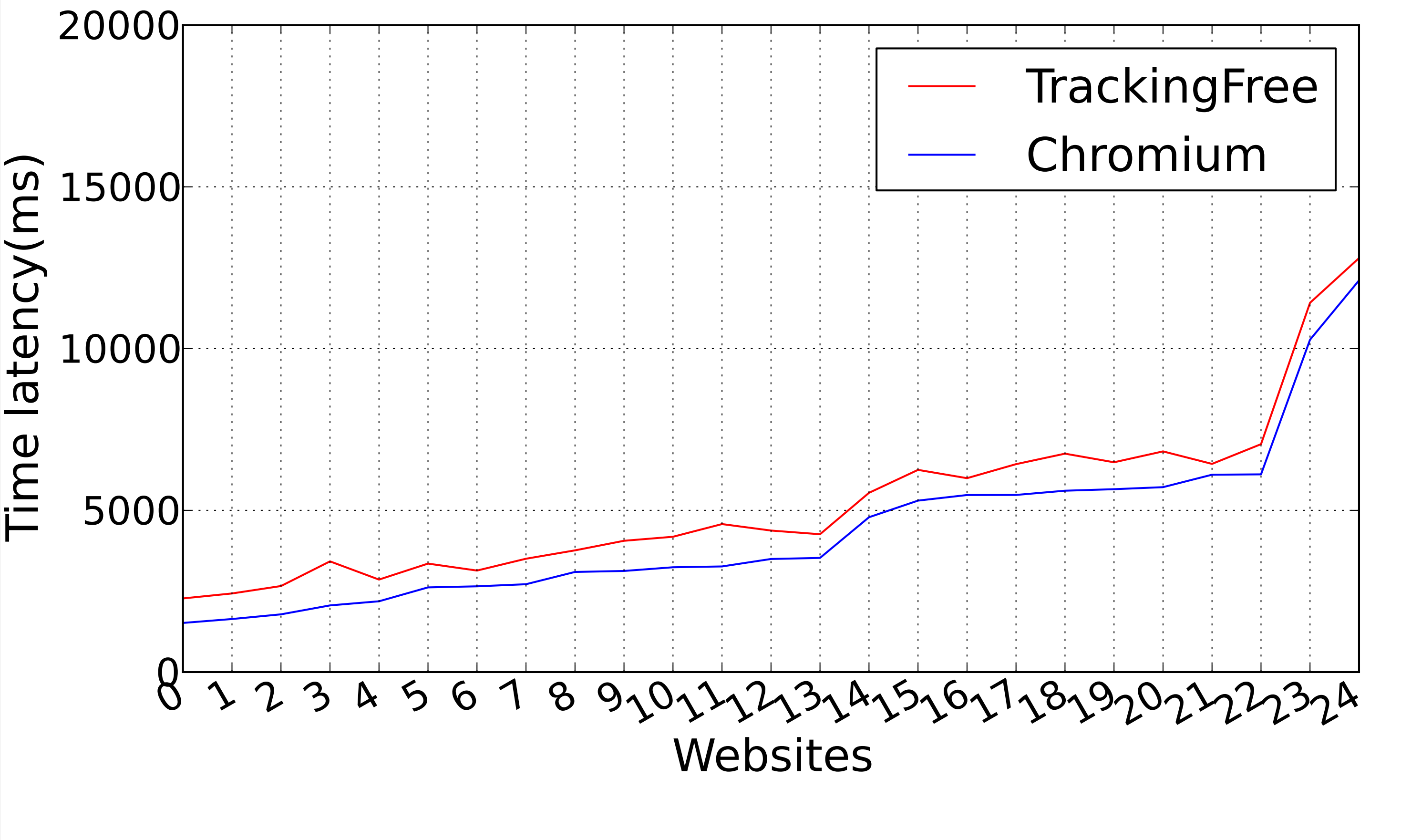
(1). Address Bar Navigation without Principal
Avg. Overhead 8.29%

Performance



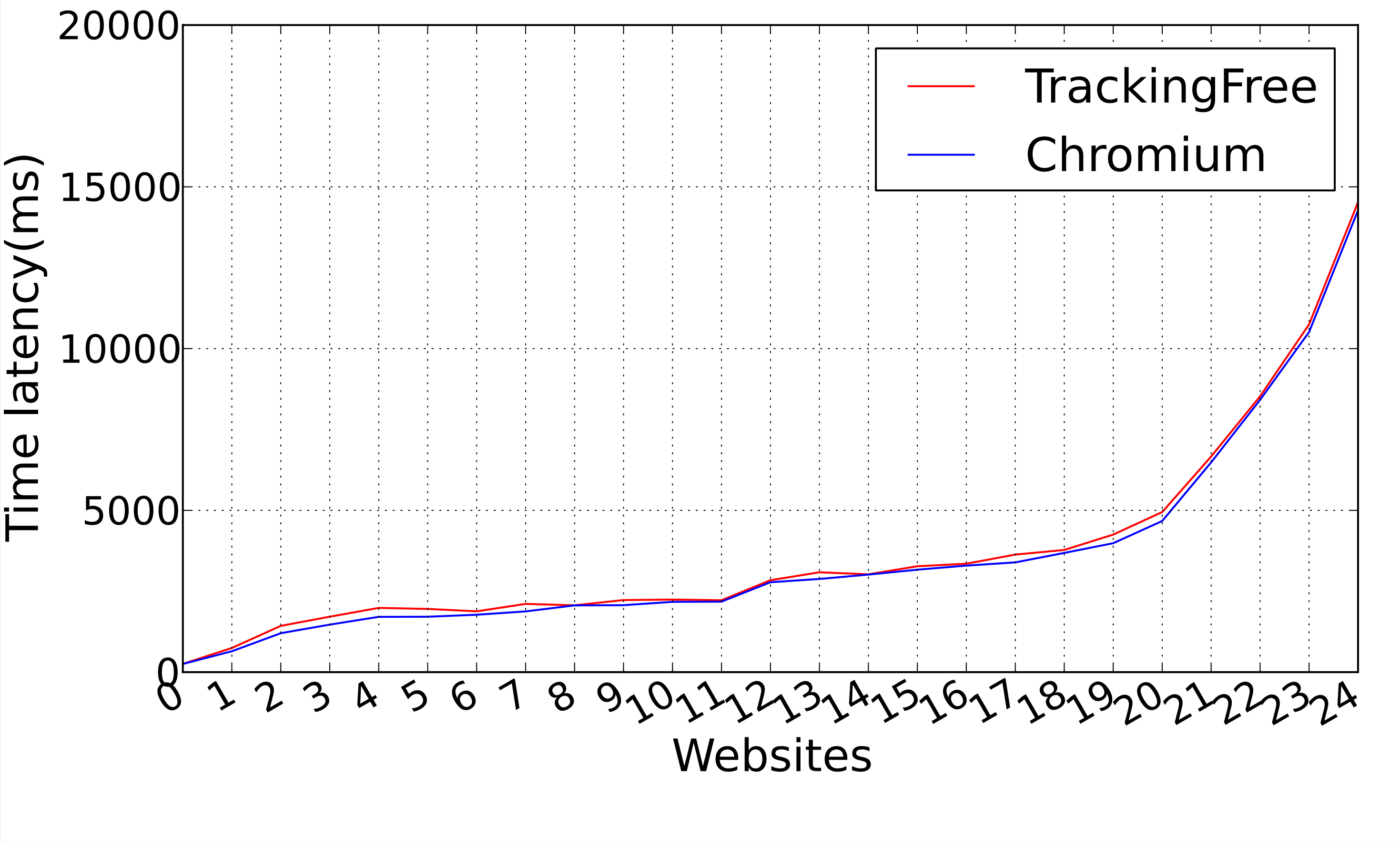
(2). Address Bar Navigation with Principal
Avg. Overhead 3.36%

Performance



(3). Cross Site Navigation
Avg. Overhead 19.43%

Performance



(4). Within-site Navigation
Avg. Overhead 4.70%

Performance

Latency Overhead Source	Cost(ms)
Principal Construction	322.36
Extra IPC	349.06
Render/JS Engine Instrumentation	139.21

Overall Overhead: ~3% - ~20%

Memory/Disk Overhead

Memory Overhead on 12 Web Pages (~25MB/Principal)

Memory	Chromium	TrackingFree	Increase
1 Principal	477.1(MB)	505(MB)	27.9(MB)
4 Principals	623.6(MB)	702.8(MB)	79.2(MB)
12 Principals	434.6(MB)	642.5(MB)	297.9(MB)

Disk Overhead on 12 Web Pages (~0.6MB/Principal)

Memory	Chromium	TrackingFree	Increase
1 Principal	21.3(MB)	21.8(MB)	0.5(MB)
4 Principals	22.5(MB)	25.9MB)	3.4(MB)
12 Principals	23.7(MB)	29.4(MB)	5.7(MB)